

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2005, 17:39:11 ; Search time 6.21033 Seconds
(without alignments)
373.662 Million cell updates/sec

Title: US-10-054-873-3
Perfect score: 33
Sequence: 1 LGTGPR 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	33	100.0	6	2	AAY42857	Aay42857 Cleavable
2	33	100.0	15	7	ADM08409	Adm08409 Canine im
3	33	100.0	15	7	ADM08322	Adm08322 Canine im
4	33	100.0	107	2	AAY42860	Aay42860 hGH-mini-
5	33	100.0	135	4	AAB93957	Aab93957 Human pro
6	33	100.0	135	7	ADJ69719	Adj69719 Human hea
7	33	100.0	150	2	AAY42861	Aay42861 Chimeric
8	33	100.0	275	8	ADS23786	Ads23786 Bacterial
9	33	100.0	388	7	ABO69171	Abo69171 Pseudomon

10	33	100.0	392	4	AAM42128	Aam42128	Human	pol
11	33	100.0	405	4	ABG22253	Abg22253	Novel	hum
12	33	100.0	585	4	AAM40342	Aam40342	Human	pol
13	33	100.0	896	8	ADR09755	Adr09755	Human	pro
14	33	100.0	952	7	ADM04599	Adm04599	Human	pro
15	33	100.0	1547	6	ABU96680	Abu96680	Human	nuc
16	33	100.0	1562	5	ABP43703	Abp43703	Alpha	-NA
17	33	100.0	2013	5	ABB47751	Abb47751	Listeria	
18	33	100.0	2013	6	ABU32684	Abu32684	Protein	e
19	33	100.0	7257	3	AAU58576	Aay58576	Sorangium	
20	31	93.9	238	4	AAU27512	Aau27512	Human	G-P
21	31	93.9	297	8	ADR43257	Adr43257	IPT'-like	
22	31	93.9	347	8	ADO29145	Ado29145	Mouse	nov
23	31	93.9	448	5	ABP66201	Abp66201	Bifidobac	
24	30	90.9	24	4	ABB39575	Abb39575	Peptide	#
25	30	90.9	24	4	AAM33134	Aam33134	Peptide	#
26	30	90.9	24	4	ABB24291	Abb24291	Protein	#
27	30	90.9	24	4	AAM72903	Aam72903	Human	bon
28	30	90.9	24	4	AAM60274	Aam60274	Human	bra
29	30	90.9	24	4	ABG54611	Abg54611	Human	liv
30	30	90.9	24	5	ABG42737	Abg42737	Human	pep
31	30	90.9	100	4	AAU49418	Aau49418	Propionib	
32	30	90.9	100	6	ABM45937	Abm45937	Propionib	
33	30	90.9	108	3	AAG04045	Aag04045	Human	sec
34	30	90.9	110	4	AAO02964	Aao02964	Human	pol
35	30	90.9	114	4	AAG89914	Aag89914	C glutami	
36	30	90.9	203	4	AAB79780	Aab79780	Corynebac	
37	30	90.9	209	4	AAB94054	Aab94054	Human	pro
38	30	90.9	209	6	ADA03124	Ada03124	Nucleotid	
39	30	90.9	222	4	AAU68588	Aau68588	Human	nov
40	30	90.9	231	6	ABO00555	Abo00555	Novel	hum
41	30	90.9	288	4	AAB67569	Aab67569	Amino	aci
42	30	90.9	288	6	ABO07232	Abo07232	Human	p53
43	30	90.9	288	8	ADP55441	Adp55441	Human	PRO
44	30	90.9	298	2	AAU73880	Aay73880	Human	pro
45	30	90.9	298	3	AAB43779	Aab43779	Human	can

ALIGNMENTS

RESULT 1

AAU42857

ID AAY42857 standard; peptide; 6 AA.

XX

AC AAY42857;

XX

DT 19-JAN-2000 (first entry)

XX

DE Cleavable peptide linker for hGH-mini-proinsulin chimeric protein.

XX

KW Linker; growth hormone; chaperone; intramolecular; insulin; precursor;

KW folding; conformation; chimeric protein; cleavable; recombinant;

KW production; yield.

XX

OS Synthetic.

XX

PN WO9950302-A1.
 XX
 PD 07-OCT-1999.
 XX
 PF 31-MAR-1998; 98WO-CN000052.
 XX
 PR 31-MAR-1998; 98WO-CN000052.
 XX
 PA (TONG-) TONGHUA GANTECH BIOTECHNOLOGY LTD.
 XX
 PI Gan Z;
 XX
 DR WPI; 1999-610839/52.
 XX
 PT New chimeric proteins containing human growth hormone fragment, used particularly for the production of human insulin.
 PT
 XX
 PS Claim 6; Page 29; 46pp; English.
 XX
 CC This sequence represents a cleavable peptide linker which is a component of the chimeric proteins hGH-mini-proinsulin (AAY42860) and the chimeric protein given in AAY42861. These chimeric proteins additionally contain an N-terminal fragment of human growth hormone (hGH) and a human insulin precursor (AAY42859) The hGH portion of the chimeric protein acts as an intramolecular chaperone (IMC) for the insulin precursor, enabling it to fold correctly. The cleavable peptide linker has a C-terminal Arg residue (AAY42857) which enables the hGH portion of the chimeric protein to be removed after folding has taken place. Production of recombinant human insulin via an hGH-proinsulin chimeric protein can provide human insulin with correctly linked cysteine bridges with fewer necessary procedural steps, and hence resulting in a higher yield of human insulin. The IMC sequences not only protect insulin sequences from intracellular degradation by a microorganism host, but also promote the folding of the fused insulin precursor, facilitate the solubility of the fusion protein and decrease the intermolecular interactions among the fusion proteins, thus allowing folding of the fused insulin precursor at commercially useful high concentrations. The procedural steps of cyanogen bromide cleavage, oxidative sulphytolysis and related purification steps can thus be eliminated, along with the use of high concentrations of mercaptan or the use of hydrophobic absorbent resins
 CC
 XX
 SQ Sequence 6 AA;

Query Match 100.0%; Score 33; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGPR 6
 |||||
 Db 1 LGTGPR 6

RESULT 2
 ADM08409
 ID ADM08409 standard; peptide; 15 AA.
 XX
 AC ADM08409;

XX
 DT 20-MAY-2004 (first entry)
 XX
 DE Canine immunoglobulin group 3 VL species framework 2 peptide 22.
 XX
 KW canine; dog; heavy; immunoglobulin; antibody light chain variable domain;
 KW antiallergic; allergy; IgE; gene therapy; group 3 species; VL framework;
 KW FR2.
 XX
 OS Canis familiaris.
 XX
 PN WO2003060080-A2.
 XX
 PD 24-JUL-2003.
 XX
 PF 20-DEC-2002; 2002WO-US041362.
 XX
 PR 21-DEC-2001; 2001US-0344874P.
 XX
 PA (IDEX-) IDEXX LAB INC.
 XX
 PI Krah ER, Guo H, Aiyappa A, Lawton R;
 XX
 DR WPI; 2003-598521/56.
 XX
 PT New canine heavy and light chain variable domain polypeptides, useful for
 PT treating canine allergy.
 XX
 PS Claim 40; Page 107; 130pp; English.
 XX
 CC The invention relates to a novel canine heavy or light chain variable
 CC domain polypeptide. The protein of the invention demonstrates
 CC antiallergic activity and may be useful for treating canine allergy,
 CC possibly via gene therapy. The current sequence is that of a canine
 CC immunoglobulin light chain variable domain framework (FR) peptide of the
 CC invention.
 XX
 SQ Sequence 15 AA;

Query Match 100.0%; Score 33; DB 7; Length 15;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
 |||||
 Db 6 LGTGPR 11

RESULT 3
 ADM08322

ID ADM08322 standard; peptide; 15 AA.
 XX
 AC ADM08322;
 XX
 DT 20-MAY-2004 (first entry)
 XX
 DE Canine immunoglobulin group 3 VL subgenus framework 2 peptide 8.

XX
 KW canine; dog; heavy; immunoglobulin; antibody light chain variable domain;
 KW antiallergic; allergy; IgE; gene therapy; group 3 subgenus; VL framework;
 KW FR2.
 XX
 OS Canis familiaris.
 XX
 PN WO2003060080-A2.
 XX
 PD 24-JUL-2003.
 XX
 PF 20-DEC-2002; 2002WO-US041362.
 XX
 PR 21-DEC-2001; 2001US-0344874P.
 XX
 PA (IDEX-) IDEXX LAB INC.
 XX
 PI Krah ER, Guo H, Aiyappa A, Lawton R;
 XX
 DR WPI; 2003-598521/56.
 XX
 PT New canine heavy and light chain variable domain polypeptides, useful for
 PT treating canine allergy.
 XX
 PS Claim 39; Page 106; 130pp; English.
 XX
 CC The invention relates to a novel canine heavy or light chain variable
 CC domain polypeptide. The protein of the invention demonstrates
 CC antiallergic activity and may be useful for treating canine allergy,
 CC possibly via gene therapy. The current sequence is that of a canine
 CC immunoglobulin light chain variable domain framework (FR) peptide of the
 CC invention.
 XX
 SQ Sequence 15 AA;

Query Match 100.0%; Score 33; DB 7; Length 15;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
 |||||
 Db 6 LGTGPR 11

RESULT 4

AAY42860

ID AAY42860 standard; protein; 107 AA.

XX

AC AAY42860;

XX

DT 19-JAN-2000 (first entry)

XX

DE hGH-mini-proinsulin chimeric protein.

XX

KW Insulin; precursor; growth hormone; chaperone; intramolecular; folding;

KW conformation; chimeric protein; cleavable; recombinant; production;

KW yield.

XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO9950302-A1.
 XX
 PD 07-OCT-1999.
 XX
 PF 31-MAR-1998; 98WO-CN000052.
 XX
 PR 31-MAR-1998; 98WO-CN000052.
 XX
 PA (TONG-) TONGHUA GANTECH BIOTECHNOLOGY LTD.
 XX
 PI Gan Z;
 XX
 DR WPI; 1999-610839/52.
 XX
 PT New chimeric proteins containing human growth hormone fragment, used
 PT particularly for the production of human insulin.
 XX
 PS Claim 13; Page 30; 46pp; English.
 XX
 CC This sequence represents a chimeric protein, hGH-mini-proinsulin. This
 CC chimeric protein contains an N-terminal fragment of human growth hormone
 CC (hGH) of the sequence given in AAY42855, a cleavable peptide linker
 CC (AAY42857), and a human insulin precursor comprising insulin A and B
 CC chains (AAY42859). The hGH portion of the chimeric protein acts as an
 CC intramolecular chaperone (IMC) for the insulin precursor, enabling it to
 CC fold correctly. The cleavable peptide linker has a C-terminal Arg residue
 CC which enables the hGH portion of the chimeric protein to be removed after
 CC folding has taken place. Production of recombinant human insulin via an
 CC hGH-proinsulin chimeric protein can provide human insulin with correctly
 CC linked cysteine bridges with fewer necessary procedural steps, and hence
 CC resulting in a higher yield of human insulin. The IMC sequences not only
 CC protect insulin sequences from intracellular degradation by a
 CC microorganism host, but also promote the folding of the fused insulin
 CC precursor, facilitate the solubility of the fusion protein and decrease
 CC the intermolecular interactions among the fusion proteins, thus allowing
 CC folding of the fused insulin precursor at commercially useful high
 CC concentrations. The procedural steps of cyanogen bromide cleavage,
 CC oxidative sulfitolysis and related purification steps can thus be
 CC eliminated, along with the use of high concentrations of mercaptan or the
 CC use of hydrophobic absorbent resins
 XX
 SQ Sequence 107 AA;

Query Match 100.0%; Score 33; DB 2; Length 107;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
 |||||
 Db 50 LGTGPR 55

RESULT 5

AAB93957

ID AAB93957 standard; protein; 135 AA.

XX

AC AAB93957;

XX

DT 26-JUN-2001 (first entry)

XX

DE Human protein sequence SEQ ID NO:14002.

XX

KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX

OS Homo sapiens.

XX

PN EP1074617-A2.

XX

PD 07-FEB-2001.

XX

PF 28-JUL-2000; 2000EP-00116126.

XX

PR 29-JUL-1999; 99JP-00248036.

PR 27-AUG-1999; 99JP-00300253.

PR 11-JAN-2000; 2000JP-00118776.

PR 02-MAY-2000; 2000JP-00183767.

PR 09-JUN-2000; 2000JP-00241899.

XX

PA (HELI-) HELIX RES INST.

XX

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX

DR WPI; 2001-318749/34.

XX

PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

XX

PS Claim 8; SEQ ID NO 14002; 2537pp + Sequence Listing; English.

XX

CC The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and

CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 135 AA;

Query Match 100.0%; Score 33; DB 4; Length 135;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
| | | | |
Db 5 LGTGPR 10

RESULT 6
ADJ69719

ID ADJ69719 standard; protein; 135 AA.
XX
AC ADJ69719;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human heat mitochondrial protein as a therapeutic target SeqID1525.
XX
KW mitochondrial; human; screening assay; diabetes mellitus;
KW Huntington's disease; osteoarthritis;
KW Leber's hereditary optic neuropathy; LHON;
KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
KW neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;
KW osteopathic; ophthalmological; cytostatic.
XX
OS Homo sapiens.
XX
PN WO2003087768-A2.
XX
PD 23-OCT-2003.
XX
PF 04-APR-2003; 2003WO-US010870.
XX
PR 12-APR-2002; 2002US-0372843P.
PR 17-JUN-2002; 2002US-0389987P.
PR 20-SEP-2002; 2002US-0412418P.
XX
PA (MITO-) MITOKOR.
PA (BUCK-) BUCK INST AGE RES.
XX
PI Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;
PI Warnock DE;
XX
DR WPI; 2003-845369/78.
XX
PT Identifying a mitochondrial target for drug screening assays and for
PT treating diseases associated with altered mitochondrial function,
PT comprises detecting a modified polypeptide in a sample and correlating

PT with the disease.

XX

PS Claim 1; SEQ ID NO 1525; 180pp; English.

XX

CC This invention relates to novel mitochondrial targets that can be used
CC for therapeutic intervention in treating a disease associated with
CC altered mitochondrial function. Specifically, it refers to a method for
CC identifying proteins of the human heart mitochondrial proteome that are
CC useful for drug screening assays, as well as therapeutic targets. The
CC present invention describes a method for identifying such proteins that
CC can be used in the treatment of various diseases associated with altered
CC mitochondrial function including diabetes mellitus, Huntington's disease,
CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
CC compositions have neuroprotective, nootropic, antidiabetic,
CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and
CC cytostatic activities. This polypeptide sequence is a human heart
CC mitochondrial protein of the invention.

XX

SQ Sequence 135 AA;

Query Match 100.0%; Score 33; DB 7; Length 135;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6

|||||

Db 5 LGTGPR 10

RESULT 7

AAY42861

ID AAY42861 standard; protein; 150 AA.

XX

AC AAY42861;

XX

DT 19-JAN-2000 (first entry)

XX

DE Chimeric protein, SEQ ID 7.

XX

KW Insulin; precursor; growth hormone; chaperone; intramolecular; folding;
KW conformation; chimeric protein; cleavable; recombinant; production;
KW yield.

XX

OS Synthetic.

OS Homo sapiens.

XX

PN WO9950302-A1.

XX

PD 07-OCT-1999.

XX

PF 31-MAR-1998; 98WO-CN000052.

XX

PR 31-MAR-1998; 98WO-CN000052.

XX

PA (TONG-) TONGHUA GANTECH BIOTECHNOLOGY LTD.

XX
 PI Gan Z;
 XX
 DR WPI; 1999-610839/52.
 XX
 PT New chimeric proteins containing human growth hormone fragment, used
 PT particularly for the production of human insulin.
 XX
 PS Claim 14; Page 30-31; 46pp; English.
 XX
 CC This sequence represents a chimeric protein, which contains an N-terminal
 CC fragment of human growth hormone (hGH) of the sequence given in AAY42856,
 CC a cleavable peptide linker (AAY42857), and a human insulin precursor
 CC comprising insulin A and B chains (AAY42859). The hGH portion of the
 CC chimeric protein acts as an intramolecular chaperone (IMC) for the
 CC insulin precursor, enabling it to fold correctly. The cleavable peptide
 CC linker has a C-terminal Arg residue which enables the hGH portion of the
 CC chimeric protein to be removed after folding has taken place. Production
 CC of recombinant human insulin via an hGH-proinsulin chimeric protein can
 CC provide human insulin with correctly linked cysteine bridges with fewer
 CC necessary procedural steps, and hence resulting in a higher yield of
 CC human insulin. The IMC sequences not only protect insulin sequences from
 CC intracellular degradation by a microorganism host, but also promote the
 CC folding of the fused insulin precursor, facilitate the solubility of the
 CC fusion protein and decrease the intermolecular interactions among the
 CC fusion proteins, thus allowing folding of the fused insulin precursor at
 CC commercially useful high concentrations. The procedural steps of cyanogen
 CC bromide cleavage, oxidative sulphytolysis and related purification steps
 CC can thus be eliminated, along with the use of high concentrations of
 CC mercaptan or the use of hydrophobic absorbent resins
 XX
 SQ Sequence 150 AA;

Query Match 100.0%; Score 33; DB 2; Length 150;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
 |||||
 Db 93 LGTGPR 98

RESULT 8

ADS23786

ID ADS23786 standard; protein; 275 AA.

XX

AC ADS23786;

XX

DT 02-DEC-2004 (first entry)

XX

DE Bacterial polypeptide #12819.

XX

KW Recombinant DNA construct; transformed plant; improved plant property;

KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;

KW pathogen tolerance; pest tolerance; plant disease resistance;

KW cell cycle pathway modification; plant growth regulator;

KW homologous recombination; seed oil yield; protein yield; carbohydrate;

KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polypeptide.
 XX
 OS Bacteria.
 XX
 PN US2003233675-A1.
 XX
 PD 18-DEC-2003.
 XX
 PF 20-FEB-2003; 2003US-00369493.
 XX
 PR 21-FEB-2002; 2002US-0360039P.
 XX
 PA (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 XX
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX
 DR WPI; 2004-061375/06.
 XX
 PT New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX
 PS Claim 1; SEQ ID NO 12819; 122pp; English.
 XX
 CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX
 SQ Sequence 275 AA;

Query Match

100.0%; Score 33; DB 8; Length 275;

Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
|||||
Db 15 LGTGPR 20

RESULT 9

ABO69171

ID ABO69171 standard; protein; 388 AA.

XX

AC ABO69171;

XX

DT 29-JUL-2004 (first entry)

XX

DE Pseudomonas aeruginosa polypeptide #1346.

XX

KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

XX

OS Pseudomonas aeruginosa.

XX

PN US6551795-B1.

XX

PD 22-APR-2003.

XX

PF 18-FEB-1999; 99US-00252991.

XX

PR 18-FEB-1998; 98US-0074788P.

PR 27-JUL-1998; 98US-0094190P.

XX

PA (GENO-) GENOME THERAPEUTICS CORP.

XX

PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX

DR WPI; 2003-615309/58.

DR N-PSDB; ABD02742.

XX

PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.

XX

PS Disclosure; SEQ ID NO 17917; 455pp; English.

XX

CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The

CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html

XX

SQ Sequence 388 AA;

Query Match 100.0%; Score 33; DB 7; Length 388;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
| | | | |
Db 179 LGTGPR 184

RESULT 10

AAM42128

ID AAM42128 standard; protein; 392 AA.

XX

AC AAM42128;

XX

DT 22-OCT-2001 (first entry)

XX

DE Human polypeptide SEQ ID NO 7059.

XX

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.

XX

OS Homo sapiens.

XX

PN WO200153312-A1.

XX

PD 26-JUL-2001.

XX

PF 26-DEC-2000; 2000WO-US034263.

XX

PR 23-DEC-1999; 99US-00471275.

PR 21-JAN-2000; 2000US-00488725.

PR 25-APR-2000; 2000US-00552317.

PR 20-JUN-2000; 2000US-00598042.

PR 19-JUL-2000; 2000US-00620312.

PR 03-AUG-2000; 2000US-00653450.

PR 14-SEP-2000; 2000US-00662191.

PR 19-OCT-2000; 2000US-00693036.

PR 29-NOV-2000; 2000US-00727344.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;

XX

DR WPI; 2001-442253/47.

DR N-PSDB; AAI61284.

XX

PT Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.

XX

PS Example 2; SEQ ID NO 7059; 10078pp; English.

XX

CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification

XX

SQ Sequence 392 AA;

Query Match 100.0%; Score 33; DB 4; Length 392;

Best Local Similarity 100.0%; Pred. No. 3.8e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6

|||||

Db 17 LGTGPR 22

RESULT 11

ABG22253

ID ABG22253 standard; protein; 405 AA.

XX

AC ABG22253;

XX

DT 18-FEB-2002 (first entry)

XX

DE Novel human diagnostic protein #22244.

XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

XX

OS Homo sapiens.

XX

PN WO200175067-A2.

XX

PD 11-OCT-2001.

XX

PF 30-MAR-2001; 2001WO-US008631.

XX

PR 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS86440.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 52612; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 405 AA;

Query Match 100.0%; Score 33; DB 4; Length 405;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
| | | | |
Db 87 LGTGPR 92

RESULT 12
AAM40342
ID AAM40342 standard; protein; 585 AA.
XX
AC AAM40342;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 3487.

XX
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US034263.
 XX
 PR 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-00552317.
 PR 20-JUN-2000; 2000US-00598042.
 PR 19-JUL-2000; 2000US-00620312.
 PR 03-AUG-2000; 2000US-00653450.
 PR 14-SEP-2000; 2000US-00662191.
 PR 19-OCT-2000; 2000US-00693036.
 PR 29-NOV-2000; 2000US-00727344.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou P, Goodrich R, Drmanac RT;
 XX
 DR WPI; 2001-442253/47.
 DR N-PSDB; AAI59498.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.
 XX
 PS Example 6; SEQ ID NO 3487; 10078pp; English.
 XX
 CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the
 CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification
 XX
 SQ Sequence 585 AA;

Query Match 100.0%; Score 33; DB 4; Length 585;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
|||||
Db 5 LGTGPR 10

RESULT 13

ADR09755

ID ADR09755 standard; protein; 896 AA.

XX

AC ADR09755;

XX

DT 04-NOV-2004 (first entry)

XX

DE Human protein useful for treating neurological disease Seq 3261.

XX

KW human; oligo-capping method; diagnostic marker; gene therapy;

KW osteoporosis; neurological disease; Alzheimer's disease;

KW Parkinson's disease; dementia; short memory; cancer;

KW sense or motor function; emotional reaction; fear response; panic;

KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;

KW tranquiliser.

XX

OS Homo sapiens.

XX

PN EP1447413-A2.

XX

PD 18-AUG-2004.

XX

PF 12-FEB-2004; 2004EP-00003145.

XX

PR 14-FEB-2003; 2003JP-00102207.

PR 09-MAY-2003; 2003JP-00131452.

XX

PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX

PI Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;

PI Wakamatsu A, Ishii S, Nagai K, Irie R;

XX

DR WPI; 2004-583265/57.

DR N-PSDB; ADR07799.

XX

PT New 1995 cDNA, useful for treating osteoporosis, neurological diseases,

PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.

XX

PS Claim 1; SEQ ID NO 3261; 2686pp; English.

XX

CC This invention relates to novel, isolated full length human cDNA

CC molecules and the encoded proteins thereof. Specifically, it refers to

CC cDNA clones obtained by an oligo-capping method, where none of these

CC clones are identical to any known human mRNAs. The present invention

CC describes an immunoassay to identify agonists and antagonists, as well as

CC antibodies, antisense molecules and siRNAs that can all be used to bind

CC to and modulate expression of the cDNA molecules. As such, these
CC molecules are useful for diagnostic markers or therapeutic targets for
CC the various diseases or morbid states. In particular, they are useful in
CC gene therapy for treating osteoporosis, neurological disease, Alzheimer's
CC disease, Parkinson's disease, dementia, short memory and various cancers,
CC as well as for maintaining equilibrium of sense or motor function, and
CC for treating emotional reaction, fear response and panic. Accordingly,
CC they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,
CC cytostatic and tranquiliser activities. This polypeptide is a protein
CC encoded by a full length human cDNA sequence of the invention. NOTE: This
CC sequence is not given in the sequence listing of the specification but
CC can be obtained on CD-ROM from the European Patent Office, Vienna Sub-
CC office.

XX

SQ Sequence 896 AA;

Query Match 100.0%; Score 33; DB 8; Length 896;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
| | | | |
Db 578 LGTGPR 583

RESULT 14

ADM04599

ID ADM04599 standard; protein; 952 AA.

XX

AC ADM04599;

XX

DT 20-MAY-2004 (first entry)

XX

DE Human protein of the invention SEQ ID NO:3284.

XX

KW human; gene therapy; diagnostic marker; pharmaceutical.

XX

OS Homo sapiens.

XX

PN EP1347046-A1.

XX

PD 24-SEP-2003.

XX

PF 12-APR-2002; 2002EP-00008400.

XX

PR 22-MAR-2002; 2002JP-00137785.

XX

PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX

PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;

PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;

PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;

XX

DR WPI; 2003-723558/69.

DR N-PSDB; ADM02156.

XX

PT New polynucleotides and polypeptides are useful in gene therapy, for

PT developing a diagnostic marker or medicines for regulating their
PT expression and activity, or as a target of gene therapy.

XX

PS Claim 1; SEQ ID NO 3284; 305pp; English.

XX

CC The invention relates to a novel human polynucleotide and the encoded
CC polypeptide. A polynucleotide of the invention may have a use in gene
CC therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful
CC as a primer for synthesizing the polynucleotide or as a probe for
CC detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are
CC useful in gene therapy, for developing a diagnostic marker or medicines
CC for regulating their expression and activity, or as a target of gene
CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
CC are useful as pharmaceutical agents. The present sequence represents a
CC protein sequence of the invention.

XX

SQ Sequence 952 AA;

Query Match 100.0%; Score 33; DB 7; Length 952;

Best Local Similarity 100.0%; Pred. No. 9.1e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6

|||||

Db 634 LGTGPR 639

RESULT 15

ABU96680

ID ABU96680 standard; protein; 1547 AA.

XX

AC ABU96680;

XX

DT 25-JUL-2003 (first entry)

XX

DE Human nucleic acid-associated protein (NAAP) #9.

XX

KW Human; nucleic acid-associated protein; cytostatic; antiarteriosclerotic;
KW anticonvulsant; nootropic; neuroprotective; cerebroprotective; anti-HIV;
KW antiallergic; antiinflammatory; thyromimetic; gene therapy;
KW cell proliferative disorder; cancer; atherosclerosis;
KW neurological disorder; epilepsy; Huntington's disease; stroke;
KW immune disorder; inflammatory disorder; AIDS; allergy;
KW developmental disorder; Hypothyroidism; Cushing's syndrome; infection;
KW protein-protein interaction; drug-target interaction;
KW gene expression profile.

XX

OS Homo sapiens.

XX

PN WO2003023003-A2.

XX

PD 20-MAR-2003.

XX

PF 05-SEP-2002; 2002WO-US028540.

XX

PR 07-SEP-2001; 2001US-0317792P.

PR 07-SEP-2001; 2001US-0317912P.

PR 14-SEP-2001; 2001US-0322270P.
PR 21-SEP-2001; 2001US-0324040P.
PR 28-SEP-2001; 2001US-0326732P.
PR 19-OCT-2001; 2001US-0346716P.
PR 25-JAN-2002; 2002US-0351749P.
PR 22-FEB-2002; 2002US-0359498P.

XX

PA (INCY-) INCYTE GENOMICS INC.

XX

PI Tang YT, Jackson JL, Griffin JA, Elliott VS, Forsythe IJ;
PI Becha SD, Richardson TW, Lee EA, Sprague WW, Emerling BM;
PI Thangavelu K, Warren BA, Tran UK, Yue H, Xu Y, Yue H, Li JX;
PI Hafalia AJA, Sanjanwala B, Marquis JP, Gorvad AE, Lee SY, Ison CH;
PI Baughn MR, Chawla NK, Nguyen DB, Swarnakar A, Zebarjadian Y, Shah P;
PI Thornton M, Yao MG, Khan FA, Gandhi AR, Yang J, Kable AE;
PI Burford N, Ramkumar J;

XX

DR WPI; 2003-313243/30.

DR N-PSDB; ACA98928.

XX

PT New human nucleic acid associated proteins (NAAP), useful for diagnosing,
PT treating and preventing diseases or conditions associated with the
PT aberrant NAAP expression e.g. cancer, AIDS, atherosclerosis, epilepsy, or
PT infections.

XX

PS Claim 1; Page 243-247; 345pp; English.

XX

CC The invention describes a novel human isolated nucleic acid-associated
CC polypeptide (NAAP). The polypeptides and polynucleotides are useful in
CC diagnosing, treating and preventing diseases or conditions associated
CC with the decreased expression or overexpression of NAAP, such as cell
CC proliferative (e.g. cancer, atherosclerosis), neurological (e.g.
CC epilepsy, Huntington's disease, stroke), immune/inflammatory (e.g. AIDS,
CC allergies) and developmental (e.g. Hypothyroidism, Cushing's syndrome)
CC disorders, or infections. These are also useful in assessing the effects
CC of exogenous compounds on the expression of nucleic acid and amino acid
CC sequences of NAAP. The NAAP or its fragments are useful in screening
CC compounds for effectiveness as agonist or antagonist of the polypeptides,
CC or in altering the expression of the target polynucleotide and compounds
CC that specifically bind to or modulate the activity of the polypeptide.
CC The microarray is useful in monitoring or measuring protein-protein
CC interactions, drug-target interactions, and gene expression profiles.
CC This is the amino acid sequence of a novel human nucleic acid-associated
CC protein (NAAP)

XX

SQ Sequence 1547 AA;

Query Match 100.0%; Score 33; DB 6; Length 1547;

Best Local Similarity 100.0%; Pred. No. 1.5e+03;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6

|||||

Db 1276 LGTGPR 1281

Search completed: February 11, 2005, 18:14:48

Job time : 10.2103 secs

OM protein - protein search, using sw model

Run on: February 11, 2005, 18:04:56 ; Search time 1.58303 Seconds
 (without alignments)
 282.936 Million cell updates/sec

Title: US-10-054-873-3
 Perfect score: 33
 Sequence: 1 LGTGPR 6

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : Issued_Patents_AA:*
 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	33	100.0	388	4	US-09-252-991A-17917	Sequence 17917, A
2	33	100.0	7257	3	US-09-335-409-5	Sequence 5, Appli
3	33	100.0	7257	3	US-09-568-102-5	Sequence 5, Appli
4	33	100.0	7257	3	US-09-567-969-5	Sequence 5, Appli
5	33	100.0	7257	3	US-09-568-480-5	Sequence 5, Appli
6	33	100.0	7257	3	US-09-568-486-5	Sequence 5, Appli
7	33	100.0	7257	3	US-09-568-472-5	Sequence 5, Appli
8	33	100.0	7257	3	US-09-567-899-5	Sequence 5, Appli
9	30	90.9	108	4	US-09-513-999C-8126	Sequence 8126, Ap
10	30	90.9	124	4	US-09-621-976-6823	Sequence 6823, Ap
11	30	90.9	211	4	US-09-949-016-11409	Sequence 11409, A

12	30	90.9	455	4	US-09-270-767-46061	Sequence 46061, A
13	30	90.9	542	4	US-09-252-991A-32139	Sequence 32139, A
14	30	90.9	683	4	US-09-949-016-11117	Sequence 11117, A
15	30	90.9	684	4	US-09-252-991A-24802	Sequence 24802, A
16	29	87.9	17	3	US-09-461-697-28	Sequence 28, Appl
17	29	87.9	20	3	US-08-724-466B-9	Sequence 9, Appli
18	29	87.9	20	3	US-08-882-164D-9	Sequence 9, Appli
19	29	87.9	58	3	US-08-927-219-39	Sequence 39, Appl
20	29	87.9	84	4	US-09-583-110-5115	Sequence 5115, Ap
21	29	87.9	143	4	US-09-252-991A-30955	Sequence 30955, A
22	29	87.9	153	4	US-09-621-976-4040	Sequence 4040, Ap
23	29	87.9	168	4	US-09-252-991A-18623	Sequence 18623, A
24	29	87.9	176	4	US-09-252-991A-24281	Sequence 24281, A
25	29	87.9	179	4	US-09-902-540-9767	Sequence 9767, Ap
26	29	87.9	198	4	US-09-248-796A-17378	Sequence 17378, A
27	29	87.9	217	4	US-09-252-991A-19838	Sequence 19838, A
28	29	87.9	239	4	US-09-949-016-9785	Sequence 9785, Ap
29	29	87.9	248	4	US-09-252-991A-21872	Sequence 21872, A
30	29	87.9	259	4	US-09-436-434-2	Sequence 2, Appli
31	29	87.9	272	4	US-09-252-991A-17461	Sequence 17461, A
32	29	87.9	274	4	US-09-252-991A-17147	Sequence 17147, A
33	29	87.9	293	4	US-09-252-991A-18793	Sequence 18793, A
34	29	87.9	305	4	US-09-710-279-1362	Sequence 1362, Ap
35	29	87.9	313	4	US-09-252-991A-23631	Sequence 23631, A
36	29	87.9	313	4	US-09-252-991A-23810	Sequence 23810, A
37	29	87.9	345	4	US-09-107-532A-4268	Sequence 4268, Ap
38	29	87.9	357	4	US-09-252-991A-28380	Sequence 28380, A
39	29	87.9	365	4	US-09-134-000C-4369	Sequence 4369, Ap
40	29	87.9	408	4	US-09-252-991A-20095	Sequence 20095, A
41	29	87.9	409	4	US-09-252-991A-23414	Sequence 23414, A
42	29	87.9	412	4	US-09-355-912A-5	Sequence 5, Appli
43	29	87.9	412	4	US-10-202-428-5	Sequence 5, Appli
44	29	87.9	430	4	US-09-252-991A-32661	Sequence 32661, A
45	29	87.9	460	4	US-09-198-452A-1085	Sequence 1085, Ap

ALIGNMENTS

RESULT 1

US-09-252-991A-17917

; Sequence 17917, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 17917

; LENGTH: 388
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17917

Query Match 100.0%; Score 33; DB 4; Length 388;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
|||||
Db 179 LGTGPR 184

RESULT 2

US-09-335-409-5
; Sequence 5, Application US/09335409
; Patent No. 6121029
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/335,409
; CURRENT FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 7257
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-335-409-5

Query Match 100.0%; Score 33; DB 3; Length 7257;
Best Local Similarity 100.0%; Pred. No. 2.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
|||||
Db 1041 LGTGPR 1046

RESULT 3

US-09-568-102-5
; Sequence 5, Application US/09568102
; Patent No. 6346404
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern

; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/568,102
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 7257
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-568-102-5

Query Match 100.0%; Score 33; DB 3; Length 7257;
Best Local Similarity 100.0%; Pred. No. 2.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
| | | | |
Db 1041 LGTGPR 1046

RESULT 4

US-09-567-969-5
; Sequence 5, Application US/09567969
; Patent No. 6355457
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/567,969
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 7257
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-567-969-5

Query Match 100.0%; Score 33; DB 3; Length 7257;
Best Local Similarity 100.0%; Pred. No. 2.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
| | | | |
Db 1041 LGTGPR 1046

RESULT 5
 US-09-568-480-5
 ; Sequence 5, Application US/09568480
 ; Patent No. 6355458
 ; GENERAL INFORMATION:
 ; APPLICANT: Schupp, Thomas
 ; APPLICANT: Ligon, James
 ; APPLICANT: Molnar, Istvan
 ; APPLICANT: Zirkle, Ross
 ; APPLICANT: Cyr, Devon
 ; APPLICANT: Goerlach, Joern
 ; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
 ; FILE REFERENCE: 4-30582A
 ; CURRENT APPLICATION NUMBER: US/09/568,480
 ; CURRENT FILING DATE: 2000-05-10
 ; PRIOR APPLICATION NUMBER: 09/335,409
 ; PRIOR FILING DATE: 1999-06-17
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 5
 ; LENGTH: 7257
 ; TYPE: PRT
 ; ORGANISM: Sorangium cellulosum
 US-09-568-480-5

Query Match 100.0%; Score 33; DB 3; Length 7257;
 Best Local Similarity 100.0%; Pred. No. 2.2e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
 |||||
 Db 1041 LGTGPR 1046

RESULT 6
 US-09-568-486-5
 ; Sequence 5, Application US/09568486
 ; Patent No. 6355459
 ; GENERAL INFORMATION:
 ; APPLICANT: Schupp, Thomas
 ; APPLICANT: Ligon, James
 ; APPLICANT: Molnar, Istvan
 ; APPLICANT: Zirkle, Ross
 ; APPLICANT: Cyr, Devon
 ; APPLICANT: Goerlach, Joern
 ; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
 ; FILE REFERENCE: 4-30582A
 ; CURRENT APPLICATION NUMBER: US/09/568,486
 ; CURRENT FILING DATE: 2000-05-10
 ; PRIOR APPLICATION NUMBER: 09/335,409
 ; PRIOR FILING DATE: 1999-06-17
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 5
 ; LENGTH: 7257
 ; TYPE: PRT
 ; ORGANISM: Sorangium cellulosum

US-09-568-486-5

Query Match 100.0%; Score 33; DB 3; Length 7257;
Best Local Similarity 100.0%; Pred. No. 2.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
|||||
Db 1041 LGTGPR 1046

RESULT 7

US-09-568-472-5

; Sequence 5, Application US/09568472
; Patent No. 6358719
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/568,472
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 7257
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-568-472-5

Query Match 100.0%; Score 33; DB 3; Length 7257;
Best Local Similarity 100.0%; Pred. No. 2.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
|||||
Db 1041 LGTGPR 1046

RESULT 8

US-09-567-899-5

; Sequence 5, Application US/09567899
; Patent No. 6383787
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES

; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/567,899
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 7257
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-567-899-5

Query Match 100.0%; Score 33; DB 3; Length 7257;
Best Local Similarity 100.0%; Pred. No. 2.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
| | | | |
Db 1041 LGTGPR 1046

RESULT 9

US-09-513-999C-8126
; Sequence 8126, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 8126
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 101
; OTHER INFORMATION: Xaa=Ile or Met
US-09-513-999C-8126

Query Match 90.9%; Score 30; DB 4; Length 108;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
: | | | | |
Db 36 VGTGPR 41

RESULT 10

US-09-621-976-6823
 ; Sequence 6823, Application US/09621976
 ; Patent No. 6639063
 ; GENERAL INFORMATION:
 ; APPLICANT: Dumas Milne Edwards, J.B.
 ; APPLICANT: Jobert, S.
 ; APPLICANT: Giordano, J.Y.
 ; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
 ; FILE REFERENCE: GENSET.054PR2
 ; CURRENT APPLICATION NUMBER: US/09/621,976
 ; CURRENT FILING DATE: 2000-07-21
 ; NUMBER OF SEQ ID NOS: 19335
 ; SOFTWARE: Patent.pm
 ; SEQ ID NO 6823
 ; LENGTH: 124
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-621-976-6823

Query Match 90.9%; Score 30; DB 4; Length 124;
 Best Local Similarity 83.3%; Pred. No. 1.8e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
 :|||||
 Db 36 VGTGPR 41

RESULT 11

US-09-949-016-11409
 ; Sequence 11409, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES
 THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 11409
 ; LENGTH: 211
 ; TYPE: PRT
 ; ORGANISM: Human
 US-09-949-016-11409

Query Match 90.9%; Score 30; DB 4; Length 211;

Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
|||||:
Db 10 LGTGPK 15

RESULT 12

US-09-270-767-46061
; Sequence 46061, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46061
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-46061

Query Match 90.9%; Score 30; DB 4; Length 455;
Best Local Similarity 83.3%; Pred. No. 5.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
|||||:
Db 107 LGTGPK 112

RESULT 13

US-09-252-991A-32139
; Sequence 32139, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32139
; LENGTH: 542
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32139

Query Match 90.9%; Score 30; DB 4; Length 542;
Best Local Similarity 83.3%; Pred. No. 6.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
:|||||
Db 261 VGTGPR 266

RESULT 14

US-09-949-016-11117
; Sequence 11117, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES
THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11117
; LENGTH: 683
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11117

Query Match 90.9%; Score 30; DB 4; Length 683;
Best Local Similarity 83.3%; Pred. No. 8.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
:|||||
Db 384 LGTGPK 389

RESULT 15

US-09-252-991A-24802
; Sequence 24802, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24802
; LENGTH: 684
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24802

Query Match 90.9%; Score 30; DB 4; Length 684;
Best Local Similarity 83.3%; Pred. No. 8.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
:|||||
Db 274 VGTGPR 279

Search completed: February 11, 2005, 18:27:03
Job time : 2.58303 secs

OM protein - protein search, using sw model

Run on: February 11, 2005, 17:42:33 ; Search time 1.14022 Seconds
(without alignments)
506.306 Million cell updates/sec

Title: US-10-054-873-3
Perfect score: 33
Sequence: 1 LGTGPR 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pirl:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	%		DB	ID	Description
		Query	Match Length			
1	33	100.0	68	2	AF2809	hypothetical prote
2	33	100.0	931	2	T49710	related to glucan
3	33	100.0	2013	2	AD1129	probable peptidogl
4	33	100.0	2013	2	AI1489	probable peptidogl
5	31	93.9	188	2	C87341	conserved hypothet
6	30	90.9	288	2	T44603	hypothetical prote
7	30	90.9	329	2	C84847	probable guanylate
8	30	90.9	387	2	T50675	guanylate kinase (
9	30	90.9	391	2	T46191	guanylate kinase-l
10	30	90.9	432	2	S49980	glutamate-5-semial
11	30	90.9	811	2	A41054	fasciclin II, tran
12	30	90.9	873	2	B41054	fasciclin II PI-li
13	30	90.9	1002	2	A36691	Ca2+-transporting

14	29	87.9	85	2	G70824	hypothetical prote
15	29	87.9	90	2	B95105	conserved hypothet
16	29	87.9	100	2	B97973	hypothetical prote
17	29	87.9	136	2	T36624	hypothetical prote
18	29	87.9	169	2	A84320	hypothetical prote
19	29	87.9	212	2	T44591	hypothetical prote
20	29	87.9	219	1	TLBPT2	tail fiber protein
21	29	87.9	226	2	S27759	maturation-associa
22	29	87.9	232	1	MMBEI3	25.5K membrane pro
23	29	87.9	238	2	T40820	proline-rich prote
24	29	87.9	243	2	S27758	maturation-associa
25	29	87.9	243	2	AH3263	exsD protein [impo
26	29	87.9	256	2	T11669	protein CPRD22, dr
27	29	87.9	259	2	T37915	hypothetical prote
28	29	87.9	260	2	A36949	28.9K basic DNA-bi
29	29	87.9	302	2	T15936	hypothetical prote
30	29	87.9	325	2	T35271	probable transcrip
31	29	87.9	359	2	T35179	vanillate O-demeth
32	29	87.9	371	1	HUBPHA	hyaluronoglucosam
33	29	87.9	371	2	B39625	T-cell receptor al
34	29	87.9	397	2	A39565	lymphoid enhancer-
35	29	87.9	399	2	A39625	T-cell receptor al
36	29	87.9	412	1	A42924	[3-methyl-2-oxobut
37	29	87.9	412	2	C72548	probable dihydroli
38	29	87.9	460	2	A72009	fumarate hydratase
39	29	87.9	460	2	B86617	fumarate hydratase
40	29	87.9	461	2	E71672	fumarate hydratase
41	29	87.9	463	2	B81725	fumarate hydratase
42	29	87.9	463	2	D97826	fumarate hydratase
43	29	87.9	463	2	H71462	probable fumarate
44	29	87.9	463	2	D87510	fumarate hydratase
45	29	87.9	464	2	H83538	fumarate hydratase

ALIGNMENTS

RESULT 1

AF2809

hypothetical protein Atul896 [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont)

C;Species: *Agrobacterium tumefaciens*

C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004

C;Accession: AF2809

R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.; Kitajima, J.P.; Okura, V.K.; Almeida Jr., N.F.; Zhou, Y.; Bovee Sr., D.; Chapman, P.; Clendenning, J.; Deatherage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClelland, E.; Palmieri, A.; Raymond, C.; Rouse, G.; Saenphimmachak, C.; Wu, Z.; Gordon, D.; Eisen, J.A.; Paulsen, I.; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001

A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, B.; Liao, L.; Kim, S.; Hendrick, C.; Zhao, Z.; Dolan, M.; Tingey, S.V.; Tomb, J.; Gordon, M.P.; Olson, M.V.; Nester, E.W.

A;Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.

A;Reference number: AB2577; MUID:21608550; PMID:11743193

A;Accession: AF2809
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-68 <KUR>
A;Cross-references: UNIPROT:Q8UE65; GB:AE008688; PIDN:AAL42892.1; PID:g17740345;
GSPDB:GN00186
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atul896
A;Map position: circular chromosome

Query Match 100.0%; Score 33; DB 2; Length 68;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
|||||
Db 36 LGTGPR 41

RESULT 2

T49710
related to glucan 1,4-alpha-glucosidase [imported] - *Neurospora crassa*
N;Alternate names: protein B23L21.230
C;Species: *Neurospora crassa*
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C;Accession: T49710
R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.;
Nyakatura, G.; Mewes, H.W.; Mannhaupt, G.
submitted to the Protein Sequence Database, May 2000
A;Reference number: Z25022
A;Accession: T49710
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-931 <SCH>
A;Cross-references: UNIPROT:Q9P5K6; EMBL:AL356172; GSPDB:GN00116;
NCSP:B23L21.230
A;Experimental source: BAC clone B23L21; strain OR74A
C;Genetics:
A;Gene: NCSP:B23L21.230
A;Map position: 6
A;Introns: 503/2

Query Match 100.0%; Score 33; DB 2; Length 931;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
|||||
Db 797 LGTGPR 802

RESULT 3

AD1129
probable peptidoglycan bound protein (LPXTG motif) lmo0435 [imported] - *Listeria monocytogenes* (strain EGD-e)
C;Species: *Listeria monocytogenes*

C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
 C;Accession: AD1129
 R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Aménd, A.; Baquero, F.; Berche, P.; Bloecker, H.; Brandt, P.; Chakraborty, T.; Charbit, A.; Chetouani, F.; Couve, E.; de Daruvar, A.; Dehoux, P.; Domann, E.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Garcia-Del Portillo, F.; Garrido, P.; Gautier, L.; Goebel, W.; Gomez-Lopez, N.; Hain, T.; Hauf, J.; Jackson, D.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mata Vicente, J.; Ng, E.; Nordsiek, G.; Novella, S.; de Pablos, B.; Perez-Diaz, J.C.; Remmel, B.; Rose, M.; Rusniok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, J.; Cossart, P.
 A;Title: Comparative genomics of *Listeria* species.
 A;Reference number: AB1077; MUID:21537279; PMID:11679669
 A;Accession: AD1129
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-2013 <GLA>
 A;Cross-references: UNIPROT:Q8Y9T8; GB:NC_003210; PIDN:CAC98514.1; PID:gl6409812; GSPDB:GN00177
 A;Experimental source: strain EGD-e
 C;Genetics:
 A;Gene: lmo0435

Query Match 100.0%; Score 33; DB 2; Length 2013;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
 |||||
 Db 892 LGTGPR 897

RESULT 4

AI1489

probable peptidoglycan bound protein (LPXTG motif) lin0457 [imported] - *Listeria innocua* (strain Clip11262)

C;Species: *Listeria innocua*

C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004

C;Accession: AI1489

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, H.; Brandt, P.; Chakraborty, T.; Charbit, A.; Chetouani, F.; Couve, E.; de Daruvar, A.; Dehoux, P.; Domann, E.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Garcia-Del Portillo, F.; Garrido, P.; Gautier, L.; Goebel, W.; Gomez-Lopez, N.; Hain, T.; Hauf, J.; Jackson, D.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001

A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mata Vicente, J.; Ng, E.; Nordsiek, G.; Novella, S.; de Pablos, B.; Perez-Diaz, J.C.; Remmel, B.; Rose, M.; Rusniok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, J.; Cossart, P.

A;Title: Comparative genomics of *Listeria* species.

A;Reference number: AB1077; MUID:21537279; PMID:11679669

A;Accession: AI1489

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-2013 <GLA>
A;Cross-references: UNIPROT:Q92EK2; GB:AL592022; PIDN:CAC95689.1; PID:g16412898;
GSPDB:GN00178
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: lin0457

Query Match 100.0%; Score 33; DB 2; Length 2013;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
| | | | |
Db 892 LGTGPR 897

RESULT 5

C87341

conserved hypothetical protein CC0742 [imported] - *Caulobacter crescentus*

C;Species: *Caulobacter crescentus*

C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004

C;Accession: C87341

R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.;
Heidelberg, J.F.; Alley, M.; Ohta, N.; Maddock, J.R.; Potocka, I.; Nelson, W.C.;
Newton, A.; Stephens, C.; Phadke, N.D.; Ely, B.; Laub, M.T.; DeBoy, R.T.;
Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolonay, J.F.; Smit, J.;
Craven, M.; Khouri, H.; Shetty, J.; Berry, K.; Utterback, T.; Tran, K.; Wolf,
A.; Vamathevan, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.;
Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A;Title: Complete Genome Sequence of *Caulobacter crescentus*.

A;Reference number: A87249; MUID:21173698; PMID:11259647

A;Accession: C87341

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-188 <STO>

A;Cross-references: UNIPROT:Q9AA64; GB:AE005673; NID:g13421975; PIDN:AAK22727.1;

GSPDB:GN00148

C;Genetics:

A;Gene: CC0742

Query Match 93.9%; Score 31; DB 2; Length 188;
Best Local Similarity 83.3%; Pred. No. 43;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
: | | | | |
Db 130 MGTGPR 135

RESULT 6

T44603

hypothetical protein CGI-83 [imported] - human

C;Species: *Homo sapiens* (man)

C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004

C;Accession: T44603

R;Lin, W.C.

submitted to the EMBL Data Library, May 1999

A;Description: Comparative gene cloning: Identification of novel human genes with *Caenorhabditis elegans* proteome as template.

A;Reference number: Z22808

A;Accession: T44603

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-288 <LIN>

A;Cross-references: UNIPROT:Q9Y392; EMBL:AF151841; PIDN:AAD34078.1

C;Genetics:

A;Map position: 8

Query Match	90.9%;	Score 30;	DB 2;	Length 288;
Best Local Similarity	83.3%;	Pred. No. 1.1e+02;		
Matches	5;	Conservative	1;	Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
 :|||||
Db 36 VGTGPR 41

RESULT 7

C84847

probable guanylate kinase [imported] - *Arabidopsis thaliana*

C;Species: *Arabidopsis thaliana* (mouse-ear cress)

C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004

C;Accession: C84847

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; Mason, T.M.; Bowman, C.L.; Barnstead, M.E.; Feldblyum, T.V.; Buell, C.R.; Ketchum, K.A.; Lee, J.J.; Ronning, C.M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.J.; Gill, J.E.; Adams, M.D.; Carrera, A.J.; Creasy, T.H.; Goodman, H.M.; Somerville, C.R.; Copenhaver, G.P.; Preuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.C.

Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.

A;Reference number: A84420; MUID:20083487; PMID:10617197

A;Accession: C84847

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-329 <STO>

A;Cross-references: UNIPROT:P93757; GB:AE002093; NID:g6598818; PIDN:AAF18683.1; GSPDB:GN00139

C;Genetics:

A;Gene: At2g41880

A;Map position: 2

Query Match	90.9%;	Score 30;	DB 2;	Length 329;
Best Local Similarity	83.3%;	Pred. No. 1.2e+02;		
Matches	5;	Conservative	1;	Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
 |||||:
Db 71 LGTGPK 76

RESULT 8

T50675

guanylate kinase (EC 2.7.4.8) [imported] - *Arabidopsis thaliana*

C;Species: *Arabidopsis thaliana* (mouse-ear cress)

C;Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 09-Jul-2004

C;Accession: T50675

R;Kumar, V.; Spangenberg, O.; Konrad, M.

Eur. J. Biochem. 267, 606-615, 2000

A;Title: Cloning of the guanylate kinase homologues AGK-1 and AGK-2 from *Arabidopsis thaliana* and characterization of AGK-1.

A;Reference number: Z25173; MUID:20098538; PMID:10632732

A;Accession: T50675

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-387 <KUM>

A;Cross-references: UNIPROT:Q9M681; EMBL:AF204677; PIDN:AAF60252.1

A;Experimental source: cultivar Columbia

C;Genetics:

A;Gene: AGK-1

A;Introns: 1/3; 39/3; 65/2; 108/3; 229/3; 315/3; 331/1; 361/2

C;Keywords: phosphotransferase

Query Match	90.9%;	Score 30;	DB 2;	Length 387;
Best Local Similarity	83.3%;	Pred. No. 1.4e+02;		
Matches	5;	Conservative	1;	Mismatches 0; Indels 0; Gaps 0;

Qy	1 LGTGPR 6
	:
Db	71 LGTGPK 76

RESULT 9

T46191

guanylate kinase-like protein - *Arabidopsis thaliana*

N;Alternate names: protein T8H10.150

C;Species: *Arabidopsis thaliana* (mouse-ear cress)

C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004

C;Accession: T46191

R;Benes, V.; Rechmann, S.; Borkova, D.; Ansorge, W.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quetier, F.; Salanoubat, M.

submitted to the Protein Sequence Database, January 2000

A;Reference number: Z23014

A;Accession: T46191

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-391 <BEN>

A;Cross-references: UNIPROT:Q9SCL8; EMBL:AL133248

A;Experimental source: cultivar Columbia; BAC clone T8H10

C;Genetics:

A;Map position: 3

A;Introns: 4/1; 40/3; 66/2; 109/3; 230/3; 271/2; 316/3; 332/1; 364/2

A;Note: T8H10.150

Query Match	90.9%;	Score 30;	DB 2;	Length 391;
Best Local Similarity	83.3%;	Pred. No. 1.4e+02;		
Matches	5;	Conservative	1;	Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
|||||:
Db 72 LGTGPK 77

RESULT 10

S49980

glutamate-5-semialdehyde dehydrogenase (EC 1.2.1.41) - *Corynebacterium glutamicum* (fragment)

N;Alternate names: gamma-glutamyl phosphate reductase

C;Species: *Corynebacterium glutamicum*

C;Date: 13-Jan-1995 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004

C;Accession: S49980

R;Serebrijski, I.; Wojcik, F.; Reyes, O.; Leblon, G.
submitted to the EMBL Data Library, November 1994

A;Description: Two loci of *Corynebacterium glutamicum* ATCC17965 that complement *Escherichia coli* mutants affected in the expression of the proA gene product.

A;Reference number: S49977

A;Accession: S49980

A;Molecule type: DNA

A;Residues: 1-432 <SER>

A;Cross-references: UNIPROT:P45638; EMBL:X82929; NID:g599719; PIDN:CAA58103.1;
PID:g599721

C;Genetics:

A;Gene: proA

C;Superfamily: gamma-glutamyl phosphate reductase

C;Keywords: oxidoreductase

Query Match 90.9%; Score 30; DB 2; Length 432;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
|||||:
Db 36 LGTGPK 41

RESULT 11

A41054

fasciclin II, transmembrane splice form precursor - fruit fly (*Drosophila melanogaster*)

C;Species: *Drosophila melanogaster*

C;Date: 21-Apr-1992 #sequence_revision 21-Apr-1992 #text_change 09-Jul-2004

C;Accession: A41054

R;Grenningloh, G.; Rehm, E.J.; Goodman, C.S.
Cell 67, 45-57, 1991

A;Title: Genetic analysis of growth cone guidance in *Drosophila*: fasciclin II functions as a neuronal recognition molecule.

A;Reference number: A41054; MUID:92005695; PMID:1913818

A;Accession: A41054

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-811 <GRE>

A;Cross-references: UNIPROT:P34082; GB:M77165; NID:g157402; PID:g157403

C;Genetics:

A;Gene: FlyBase:Fas2

A;Cross-references: FlyBase:FBgn0000635

C;Superfamily: neural cell adhesion molecule; fibronectin type III repeat
homology; immunoglobulin homology
C;Keywords: membrane protein

Query Match 90.9%; Score 30; DB 2; Length 811;
Best Local Similarity 83.3%; Pred. No. 3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
:|||||
Db 481 VGTGPR 486

RESULT 12

B41054

fasciclin II PI-linked splice form precursor - fruit fly (*Drosophila melanogaster*)

C;Species: *Drosophila melanogaster*

C;Date: 21-Apr-1992 #sequence_revision 21-Apr-1992 #text_change 17-Mar-2000

C;Accession: B41054

R;Grenningloh, G.; Rehm, E.J.; Goodman, C.S.
Cell 67, 45-57, 1991

A;Title: Genetic analysis of growth cone guidance in *Drosophila*: fasciclin II
functions as a neuronal recognition molecule.

A;Reference number: A41054; MUID:92005695; PMID:1913818

A;Accession: B41054

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-873 <GRE>

A;Cross-references: GB:M77166

C;Genetics:

A;Gene: FlyBase:Fas2

A;Cross-references: FlyBase:FBgn0000635

C;Superfamily: neural cell adhesion molecule; fibronectin type III repeat
homology; immunoglobulin homology

C;Keywords: transmembrane protein

Query Match 90.9%; Score 30; DB 2; Length 873;
Best Local Similarity 83.3%; Pred. No. 3.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
:|||||
Db 481 VGTGPR 486

RESULT 13

A36691

Ca²⁺-transporting ATPase (EC 3.6.3.8), sarcoplasmic reticulum - fruit fly
(*Drosophila melanogaster*)

C;Species: *Drosophila melanogaster*

C;Date: 28-Jun-1991 #sequence_revision 30-Jan-1993 #text_change 09-Jul-2004

C;Accession: A36691; S07050

R;Magyar, A.; Varadi, A.

Biochem. Biophys. Res. Commun. 173, 872-877, 1990

A;Title: Molecular cloning and chromosomal localization of a sarco/endoplasmic
reticulum-type Ca(2+)-ATPase of *Drosophila melanogaster*.

A;Reference number: A36691; MUID:91097592; PMID:2148477
 A;Accession: A36691
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-1002 <MAG>
 A;Cross-references: UNIPROT:P22700; GB:M62892; NID:g158415; PIDN:AAB00735.1; PID:g158416
 R;Varadi, A.; Gilmore-Heber, M.; Benz Jr., E.J.
 FEBS Lett. 258, 203-207, 1989
 A;Title: Amplification of the phosphorylation site - ATP-binding site cDNA fragment of the Na(+),K(+)-ATPase and the Ca(2+)-ATPase of Drosophila melanogaster by polymerase chain reaction.
 A;Reference number: S07049; MUID:90092469; PMID:2557235
 A;Accession: S07050
 A;Molecule type: mRNA
 A;Residues: 357-513 <VAR>
 A;Cross-references: EMBL:X17472
 A;Note: the authors translated the codon CTC for residue 1 as Thr; the sequence shown follows the authors' translation
 C;Genetics:
 A;Gene: FlyBase:Ca-P60A
 A;Cross-references: FlyBase:FBgn0004551
 C;Function:
 A;Description: catalyzes active transport of Ca²⁺ ions accross cellular membranes; Ca²⁺ pump
 C;Superfamily: Na⁺/K⁺-transporting ATPase alpha chain; ATPase nucleotide-binding domain homology
 C;Keywords: ATP; calcium transport; hydrolase; phosphoprotein; transmembrane protein
 F;40-57/Domain: calcium binding #status predicted <CA1>
 F;60-78/Domain: transmembrane #status predicted <TM01>
 F;87-107/Domain: transmembrane #status predicted <TM02>
 F;108-257/Domain: intracellular #status predicted <INT1>
 F;111-131/Domain: calcium binding #status predicted <CA2>
 F;132-238/Domain: transduction #status predicted <TSD>
 F;258-277/Domain: transmembrane #status predicted <TM03>
 F;288-307/Domain: transmembrane #status predicted <TM04>
 F;308-760/Domain: intracellular #status predicted <INT2>
 F;310-329/Domain: calcium binding #status predicted <CA3>
 F;330-505/Domain: catalytic #status predicted <PHY>
 F;506-680/Domain: ATP binding #status predicted <ATP>
 F;595-768/Domain: ATPase nucleotide-binding domain homology <ATN>
 F;763-784/Domain: transmembrane #status predicted <TM05>
 F;788-809/Domain: transmembrane #status predicted <TM06>
 F;837-857/Domain: transmembrane #status predicted <TM07>
 F;894-913/Domain: transmembrane #status predicted <TM08>
 F;931-950/Domain: transmembrane #status predicted <TM09>
 F;959-980/Domain: transmembrane #status predicted <TM10>
 F;351/Active site: Asp (aspartylphosphate intermediate) #status predicted
 F;515/Binding site: ATP (Lys) #status predicted

Query Match 90.9%; Score 30; DB 2; Length 1002;
 Best Local Similarity 83.3%; Pred. No. 3.8e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
 ||||:

RESULT 14

G70824

hypothetical protein Rv0748 - Mycobacterium tuberculosis (strain H37RV)

C;Species: Mycobacterium tuberculosis

C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004

C;Accession: G70824

R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.V.; Eiglmeier, K.; Gas, S.; Barry III, C.E.; Tekaia, F.; Badcock, K.; Basham, D.; Brown, D.; Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Hornsby, T.; Jagels, K.; Krogh, A.; McLean, J.; Moule, S.; Murphy, L.; Oliver, S.; Osborne, J.; Quail, M.A.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A;Authors: Sqaes, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence.

A;Reference number: A70500; MUID:98295987; PMID:9634230

A;Accession: G70824

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-85 <COL>

A;Cross-references: UNIPROT:O53811; GB:AL021958; GB:AL123456; NID:g3261536;

PIDN:CAAL7515.1; PID:el253286; PID:g2911022

A;Experimental source: strain H37Rv

C;Genetics:

A;Gene: Rv0748

Query Match 87.9%; Score 29; DB 2; Length 85;

Best Local Similarity 100.0%; Pred. No. 49;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GTGPR 6

|||||

Db 56 GTGPR 60

RESULT 15

B95105

conserved hypothetical protein SP0910 [imported] - Streptococcus pneumoniae (strain TIGR4)

C;Species: Streptococcus pneumoniae

C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004

C;Accession: B95105

R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heidelberg, J.; DeBoy, R.T.; Haft, D.H.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.; Kolonay, J.F.; Nelson, W.C.; Peterson, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple, E.; Khouri, H.; Wolf, A.M.; Utterback, T.R.; Hansen, C.L.; McDonald, L.A.; Feldblyum, T.V.; Angiuoli, S.; Dickinson, T.; Hickey, E.K.; Holt, I.E.

Science 293, 498-506, 2001

A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison, D.A.; Hollingshead, S.K.; Fraser, C.M.

A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A;Reference number: A95000; MUID:21357209; PMID:11463916
A;Accession: B95105
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-90 <KUR>
A;Cross-references: UNIPROT:Q97RB4; GB:AE005672; PIDN:AAK75035.1; PID:g14972384; GSPDB:GN00164; TIGR:SP4SP0910
A;Experimental source: strain TIGR4
C;Genetics:
A;Gene: SP0910

Query Match 87.9%; Score 29; DB 2; Length 90;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GTGPR 6
|||||
Db 70 GTGPR 74

Search completed: February 11, 2005, 18:24:34
Job time : 4.14022 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2005, 18:23:02 ; Search time 4.36162 Seconds
(without alignments)
449.487 Million cell updates/sec

Title: US-10-054-873-3
Perfect score: 33
Sequence: 1 LGTGPR 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1376875 seqs, 326749119 residues

Total number of hits satisfying chosen parameters: 1376875

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	33	100.0	6	13	US-10-054-873-3	Sequence 3, Appli
2	33	100.0	15	16	US-10-327-598-290	Sequence 290, App
3	33	100.0	82	16	US-10-437-963-168327	Sequence 168327,
4	33	100.0	107	13	US-10-054-873-6	Sequence 6, Appli
5	33	100.0	135	16	US-10-408-765A-1525	Sequence 1525, Ap
6	33	100.0	150	13	US-10-054-873-7	Sequence 7, Appli
7	33	100.0	275	15	US-10-369-493-12819	Sequence 12819, A
8	33	100.0	802	16	US-10-437-963-178776	Sequence 178776,
9	33	100.0	952	15	US-10-108-260A-3284	Sequence 3284, Ap
10	33	100.0	2013	15	US-10-282-122A-60608	Sequence 60608, A
11	33	100.0	7257	13	US-10-014-717-5	Sequence 5, Appli
12	31	93.9	455	16	US-10-437-963-154712	Sequence 154712,
13	31	93.9	1963	16	US-10-437-963-155445	Sequence 155445,
14	30	90.9	24	9	US-09-864-761-39589	Sequence 39589, A
15	30	90.9	78	16	US-10-437-963-188463	Sequence 188463,
16	30	90.9	78	16	US-10-767-701-56946	Sequence 56946, A
17	30	90.9	88	15	US-10-424-599-183036	Sequence 183036,
18	30	90.9	114	9	US-09-738-626-3668	Sequence 3668, Ap
19	30	90.9	186	16	US-10-767-701-36431	Sequence 36431, A
20	30	90.9	208	14	US-10-233-553-30	Sequence 30, Appl
21	30	90.9	222	15	US-10-240-145-161	Sequence 161, App
22	30	90.9	231	15	US-10-243-552-564	Sequence 564, App
23	30	90.9	268	15	US-10-424-599-147656	Sequence 147656,
24	30	90.9	274	15	US-10-424-599-164540	Sequence 164540,
25	30	90.9	275	15	US-10-425-114-48504	Sequence 48504, A
26	30	90.9	298	9	US-09-925-301-1224	Sequence 1224, Ap
27	30	90.9	298	15	US-10-296-115-1307	Sequence 1307, Ap
28	30	90.9	309	15	US-10-425-114-47615	Sequence 47615, A
29	30	90.9	309	15	US-10-425-114-48066	Sequence 48066, A
30	30	90.9	377	16	US-10-437-963-173919	Sequence 173919,
31	30	90.9	432	9	US-09-738-626-6081	Sequence 6081, Ap
32	30	90.9	432	17	US-10-494-675-106	Sequence 106, App
33	30	90.9	471	14	US-10-029-386-31956	Sequence 31956, A
34	30	90.9	575	15	US-10-425-114-64569	Sequence 64569, A
35	30	90.9	575	15	US-10-425-114-68806	Sequence 68806, A
36	30	90.9	676	15	US-10-374-979-101	Sequence 101, App
37	30	90.9	676	15	US-10-182-936A-101	Sequence 101, App
38	30	90.9	804	14	US-10-128-714-3088	Sequence 3088, Ap
39	30	90.9	1028	14	US-10-128-714-8088	Sequence 8088, Ap
40	30	90.9	1148	16	US-10-437-963-141916	Sequence 141916,
41	30	90.9	1204	16	US-10-437-963-188462	Sequence 188462,
42	30	90.9	1233	16	US-10-437-963-188471	Sequence 188471,
43	30	90.9	1467	16	US-10-437-963-106949	Sequence 106949,
44	30	90.9	1540	16	US-10-437-963-188461	Sequence 188461,
45	30	90.9	2071	16	US-10-437-963-103806	Sequence 103806,

ALIGNMENTS

RESULT 1

US-10-054-873-3

; Sequence 3, Application US/10054873

; Publication No. US20020164712A1

```

; GENERAL INFORMATION:
; APPLICANT: Gan, Zhong Ru
; TITLE OF INVENTION: Chimeric Protein Containing an
; Intramolecular Chaperone-Like Sequence
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/054,873
; FILING DATE: 22-Jan-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/CN98/00052
; FILING DATE: 31-MAR-1998
; APPLICATION NUMBER: US 09/423,100
; FILING DATE: 11-DEC-2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Mycroft, Frank J
; REGISTRATION NUMBER: 46,946
; REFERENCE/DOCKET NUMBER: 020167-000130US
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-054-873-3

```

```

Query Match          100.0%; Score 33; DB 13; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.2e+06;
Matches      6; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      1 LGTGPR 6
        |||||
Db      1 LGTGPR 6

```

```

RESULT 2
US-10-327-598-290
; Sequence 290, Application US/10327598
; Publication No. US20040181039A1
; GENERAL INFORMATION:
; APPLICANT: Krah, Eugene
; APPLICANT: Guo, Honliang
; APPLICANT: Aiyappa, Ashok

```

; APPLICANT: Lawton, Robert
; TITLE OF INVENTION: Canine Immunoglobulin Variable Domains, Caninized
Antibodies, and Methods
; TITLE OF INVENTION: for Making and Using Them
; FILE REFERENCE: 01-799-A
; CURRENT APPLICATION NUMBER: US/10/327,598
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US 60/344,874
; PRIOR FILING DATE: 2001-12-21
; NUMBER OF SEQ ID NOS: 1139
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 290
; LENGTH: 15
; TYPE: PRT
; ORGANISM: canis familiaris;
US-10-327-598-290

Query Match 100.0%; Score 33; DB 16; Length 15;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
| | | | |
Db 6 LGTGPR 11

RESULT 3

US-10-437-963-168327

; Sequence 168327, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules
Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 168327
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(82)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_66852C.1.pep
US-10-437-963-168327

Query Match 100.0%; Score 33; DB 16; Length 82;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
|||||
Db 13 LGTGPR 18

RESULT 4

US-10-054-873-6

; Sequence 6, Application US/10054873

; Publication No. US20020164712A1

; GENERAL INFORMATION:

; APPLICANT: Gan, Zhong Ru

; TITLE OF INVENTION: Chimeric Protein Containing an
; Intramolecular Chaperone-Like Sequence

; NUMBER OF SEQUENCES: 7

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew LLP

; STREET: Two Embarcadero Center, Eighth Floor

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/054,873

; FILING DATE: 22-Jan-2002

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: WO PCT/CN98/00052

; FILING DATE: 31-MAR-1998

; APPLICATION NUMBER: US 09/423,100

; FILING DATE: 11-DEC-2000

; ATTORNEY/AGENT INFORMATION:

; NAME: Mycroft, Frank J

; REGISTRATION NUMBER: 46,946

; REFERENCE/DOCKET NUMBER: 020167-000130US

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 107 amino acids

; TYPE: amino acid

; STRANDEDNESS: <Unknown>

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 6:

US-10-054-873-6

Query Match 100.0%; Score 33; DB 13; Length 107;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
 |||||
Db 50 LGTGPR 55

RESULT 5

US-10-408-765A-1525

; Sequence 1525, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1525
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-1525

Query Match 100.0%; Score 33; DB 16; Length 135;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
 |||||
Db 5 LGTGPR 10

RESULT 6

US-10-054-873-7

; Sequence 7, Application US/10054873
; Publication No. US20020164712A1
; GENERAL INFORMATION:
; APPLICANT: Gan, Zhong Ru
; TITLE OF INVENTION: Chimeric Protein Containing an
; Intramolecular Chaperone-Like Sequence
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/054,873
; FILING DATE: 22-Jan-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/CN98/00052
; FILING DATE: 31-MAR-1998
; APPLICATION NUMBER: US 09/423,100
; FILING DATE: 11-DEC-2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Mycroft, Frank J
; REGISTRATION NUMBER: 46,946
; REFERENCE/DOCKET NUMBER: 020167-000130US
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 150 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-054-873-7

```

```

Query Match          100.0%;  Score 33;  DB 13;  Length 150;
Best Local Similarity 100.0%;  Pred. No. 1.8e+02;
Matches      6;  Conservative    0;  Mismatches    0;  Indels      0;  Gaps      0;

```

```

Qy      1 LGTGPR 6
        |||||
Db      93 LGTGPR 98

```

RESULT 7

```

US-10-369-493-12819
; Sequence 12819, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION
OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 12819
; LENGTH: 275

```

; TYPE: PRT
; ORGANISM: Aspergillus nidulans
US-10-369-493-12819

Query Match 100.0%; Score 33; DB 15; Length 275;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
|||||
Db 15 LGTGPR 20

RESULT 8

US-10-437-963-178776
; Sequence 178776, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules
Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 178776
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_7629C.1.pep
US-10-437-963-178776

Query Match 100.0%; Score 33; DB 16; Length 802;
Best Local Similarity 100.0%; Pred. No. 8.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
|||||
Db 279 LGTGPR 284

RESULT 9

US-10-108-260A-3284
; Sequence 3284, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA

; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3284
; LENGTH: 952
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-3284

Query Match 100.0%; Score 33; DB 15; Length 952;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
| | | | |
Db 634 LGTGPR 639

RESULT 10

US-10-282-122A-60608
; Sequence 60608, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60608
; LENGTH: 2013
; TYPE: PRT
; ORGANISM: *Listeria monocytogenes*
US-10-282-122A-60608

Query Match 100.0%; Score 33; DB 15; Length 2013;
Best Local Similarity 100.0%; Pred. No. 2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
| | | | |
Db 892 LGTGPR 897

RESULT 11

US-10-014-717-5

; Sequence 5, Application US/10014717
; Publication No. US20020192778A1
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/10/014,717
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US/09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 7257
; TYPE: PRT
; ORGANISM: *Sorangium cellulosum*
US-10-014-717-5

Query Match 100.0%; Score 33; DB 13; Length 7257;
Best Local Similarity 100.0%; Pred. No. 6.6e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
| | | | |
Db 1041 LGTGPR 1046

RESULT 12

US-10-437-963-154712
; Sequence 154712, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules
Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 154712
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(455)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_54545C.1.pep
US-10-437-963-154712

Query Match 93.9%; Score 31; DB 16; Length 455;
Best Local Similarity 83.3%; Pred. No. 1.2e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
:|||||
Db 338 MGTGPR 343

RESULT 13

US-10-437-963-155445
; Sequence 155445, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules
Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 155445
; LENGTH: 1963
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_55207C.1.pep
US-10-437-963-155445

Query Match 93.9%; Score 31; DB 16; Length 1963;
Best Local Similarity 83.3%; Pred. No. 4.6e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
:|||||
Db 336 MGTGPR 341

RESULT 14

US-09-864-761-39589

; Sequence 39589, Application US/09864761

; Patent No. US20020048763A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharron G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; APPLICANT: Chen, Wensheng

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES
USEFUL FOR

; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

; FILE REFERENCE: Aeomica-X-1

; CURRENT APPLICATION NUMBER: US/09/864,761

; CURRENT FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: US 60/180,312

; PRIOR FILING DATE: 2000-02-04

; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: US 09/632,366

; PRIOR FILING DATE: 2000-08-03

; PRIOR APPLICATION NUMBER: GB 24263.6

; PRIOR FILING DATE: 2000-10-04

; PRIOR APPLICATION NUMBER: US 60/236,359

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: PCT/US01/00666

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00667

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00664

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00669

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00665

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00668

; PRIOR FILING DATE: 2001-01-30


```
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 39589
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004061.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.84
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7
US-09-864-761-39589
```

```
Query Match          90.9%; Score 30; DB 9; Length 24;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches      5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 LGTGPR 6
        |||||:
Db      13 LGTGPK 18
```

RESULT 15

US-10-437-963-188463

```
; Sequence 188463, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules
Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
```

; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 188463
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_85064C.1.pep
US-10-437-963-188463

Query Match 90.9%; Score 30; DB 16; Length 78;
Best Local Similarity 83.3%; Pred. No. 3.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
:|||||
Db 30 VGTGPR 35

Search completed: February 11, 2005, 19:03:51
Job time : 5.36162 secs

OM protein - protein search, using sw model

Run on: February 11, 2005, 17:42:04 ; Search time 5.23616 Seconds
 (without alignments)
 586.780 Million cell updates/sec

Title: US-10-054-873-3
 Perfect score: 33
 Sequence: 1 LGTGPR 6

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : UniProt_03:*
 1: uniprot_sprot:*
 2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query			DB ID	Description
		Match	Length			
1	33	100.0	68	2	Q8UE65	Q8ue65 agrobacteri
2	33	100.0	135	2	Q9HA29	Q9ha29 homo sapien
3	33	100.0	277	2	Q8L1U5	Q8l1u5 bordetella
4	33	100.0	287	2	Q7W361	Q7w361 bordetella
5	33	100.0	287	2	Q7WEH8	Q7weh8 bordetella
6	33	100.0	289	2	Q7VSQ6	Q7vsq6 bordetella
7	33	100.0	491	2	Q8P6Z3	Q8p6z3 xanthomonas
8	33	100.0	563	2	Q8I0F1	Q8i0f1 drosophila
9	33	100.0	585	2	Q8NEG7	Q8neg7 homo sapien
10	33	100.0	736	2	Q6YBW4	Q6ybw4 oryctolagus
11	33	100.0	896	2	Q6ZRS4	Q6zrs4 homo sapien
12	33	100.0	946	2	Q9P5K6	Q9p5k6 neurospora
13	33	100.0	1095	2	Q8KLS0	Q8kls0 rhodobacter
14	33	100.0	1522	2	O15069	O15069 homo sapien
15	33	100.0	2013	2	Q8Y9T8	Q8y9t8 listeria mo

16	33	100.0	2013	2	Q92EK2	Q92ek2	listeria in
17	33	100.0	7257	2	Q9L8C7	Q9l8c7	polyangium
18	31	93.9	188	2	Q9AA64	Q9aa64	caulobacter
19	31	93.9	297	2	Q8VQS5	Q8vqs5	methylobact
20	31	93.9	376	2	Q7UWF9	Q7uwf9	rhodopirell
21	31	93.9	448	2	Q8G430	Q8g430	bifidobacte
22	31	93.9	485	2	Q6YX20	Q6yx20	oryza sativ
23	31	93.9	541	2	Q6ZFA9	Q6zfa9	oryza sativ
24	31	93.9	543	2	Q6ZFS1	Q6zfs1	oryza sativ
25	31	93.9	561	2	Q9DK04	Q9dk04	allpahuayo
26	31	93.9	579	2	Q9LD30	Q9ld30	crypthecodi
27	31	93.9	656	2	O96529	O96529	meloidogyne
28	31	93.9	656	2	Q9XYA9	Q9xya9	meloidogyne
29	31	93.9	676	2	Q6ZFB7	Q6zfb7	oryza sativ
30	31	93.9	688	2	Q6ZFT2	Q6zft2	oryza sativ
31	30	90.9	114	2	Q8NU02	Q8nu02	corynebacte
32	30	90.9	128	2	Q8FU73	Q8fu73	corynebacte
33	30	90.9	151	2	Q93WV9	Q93wv9	musa acumin
34	30	90.9	156	2	Q68E52	Q68e52	aeromonas p
35	30	90.9	196	2	Q6MPT9	Q6mpt9	bdellovibri
36	30	90.9	200	2	Q8K VW0	Q8kvw0	ruegeria sp
37	30	90.9	209	2	Q9H9X7	Q9h9x7	homo sapien
38	30	90.9	269	2	Q8LHV5	Q8lhv5	oryza sativ
39	30	90.9	288	2	Q9Y392	Q9y392	homo sapien
40	30	90.9	290	2	Q6JPQ9	Q6j pq9	uncultured
41	30	90.9	324	2	Q8P5X3	Q8p5x3	xanthomonas
42	30	90.9	325	2	Q8WTP8	Q8wtp8	homo sapien
43	30	90.9	327	2	Q9BSA5	Q9bsa5	homo sapien
44	30	90.9	387	2	P93757	P93757	arabidopsis
45	30	90.9	387	2	Q683H2	Q683h2	arabidopsis

ALIGNMENTS

RESULT 1

Q8UE65

ID Q8UE65 PRELIMINARY; PRT; 68 AA.
AC Q8UE65;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein Atul896.
GN OrderedLocusNames=Atul896;
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Dupont;
RX MEDLINE=21608550; PubMed=11743193; DOI=10.1126/science.1066804;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kuttyavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,

RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nester E.W.;
 RT "The genome of the natural genetic engineer *Agrobacterium tumefaciens*
 RT C58.";
 RL Science 294:2317-2323(2001).
 DR EMBL; AE009143; AAL42892.1; -.
 DR PIR; AF2809; AF2809.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 68 AA; 7596 MW; AE8CBD8946139A9F CRC64;

Query Match 100.0%; Score 33; DB 2; Length 68;
 Best Local Similarity 100.0%; Pred. No. 43;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGPR 6
 |||||
 Db 36 LGTGPR 41

RESULT 2

Q9HA29

ID Q9HA29 PRELIMINARY; PRT; 135 AA.
 AC Q9HA29;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein FLJ12345.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Mammary gland;
 RX PubMed=14702039; DOI=10.1038/ngl285;
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
 RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
 RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K.,
 RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
 RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
 RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
 RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
 RA Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y., Ishida S.,
 RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
 RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
 RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
 RA Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
 RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
 RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,

RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
 RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
 RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs.";
 RL Nat. Genet. 36:40-45(2004).
 DR EMBL; AK022407; BAB14030.1; -.
 SQ SEQUENCE 135 AA; 14034 MW; 0D37366C979CDDA8 CRC64;

Query Match 100.0%; Score 33; DB 2; Length 135;
 Best Local Similarity 100.0%; Pred. No. 81;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
 |||||
 Db 5 LGTGPR 10

RESULT 3

Q8L1U5

ID Q8L1U5 PRELIMINARY; PRT; 277 AA.
 AC Q8L1U5;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE BhuT.
 GN Name=bhuT;
 OS Bordetella avium.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 OX NCBI_TaxID=521;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=4169;
 RX MEDLINE=21481973; PubMed=11598070;
 RX DOI=10.1128/IAI.69.11.6951-6961.2001;
 RA Kirby A.E., Metzger D.J., Murphy E.R., Connell T.D.;
 RT "Heme utilization in Bordetella avium is regulated by RhuI, a heme-
 RT responsive extracytoplasmic function sigma factor.";
 RL Infect. Immun. 69:6951-6961(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=4169;
 RX MEDLINE=22215669; PubMed=12228263;
 RX DOI=10.1128/IAI.70.10.5390-5403.2002;
 RA Murphy E.R., Sacco R.E., Dickenson A., Metzger D.J., Hu Y.,
 RA Orndorff P.E., Connell T.D.;
 RT "BhuR, a virulence-associated outer membrane protein of Bordetella
 RT avium, is required for the acquisition of iron from heme and
 RT hemoproteins.";
 RL Infect. Immun. 70:5390-5403(2002).

DR EMBL; AY095952; AAM28270.1; -.
DR GO; GO:0005381; F:iron ion transporter activity; IEA.
DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
DR InterPro; IPR002491; Peripla_BP.
DR Pfam; PF01497; Peripla_BP_2; 1.
SQ SEQUENCE 277 AA; 28898 MW; F9CDDCCD2AA37B4D CRC64;

Query Match 100.0%; Score 33; DB 2; Length 277;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
|||||
Db 252 LGTGPR 257

RESULT 4

Q7W361

ID Q7W361 PRELIMINARY; PRT; 287 AA.
AC Q7W361;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative hemin binding protein.
GN Name=bhuT; OrderedLocusNames=BPP4187;
OS Bordetella parapertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=519;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=12822 / ATCC BAA-587;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ngl227;
RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,
RA Rabbिनowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640436; CAE39466.1; -.
DR GO; GO:0005381; F:iron ion transporter activity; IEA.
DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
DR InterPro; IPR002491; Peripla_BP.
DR Pfam; PF01497; Peripla_BP_2; 1.
KW Complete proteome.
SQ SEQUENCE 287 AA; 29393 MW; 86F8317AD5241C40 CRC64;

Query Match 100.0%; Score 33; DB 2; Length 287;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
|||||
Db 259 LGTGPR 264

RESULT 5

Q7WEH8

ID Q7WEH8 PRELIMINARY; PRT; 287 AA.
AC Q7WEH8;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative hemin binding protein.
GN Name=bhuT; OrderedLocusNames=BB4657;
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RB50 / ATCC BAA-588;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ngl227;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640451; CAE35019.1; -.
DR GO; GO:0005381; F:iron ion transporter activity; IEA.
DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
DR InterPro; IPR002491; Peripla_BP.
DR Pfam; PF01497; Peripla_BP_2; 1.
KW Complete proteome.
SQ SEQUENCE 287 AA; 29363 MW; 96F9317AC5251031 CRC64;

Query Match 100.0%; Score 33; DB 2; Length 287;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
|||||
Db 259 LGTGPR 264

RESULT 6

Q7VSQ6

ID Q7VSQ6 PRELIMINARY; PRT; 289 AA.
AC Q7VSQ6;

DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Putative hemin binding protein.
 GN Name=bhuT; OrderedLocusNames=BP0345;
 OS Bordetella pertussis.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 OX NCBI_TaxID=520;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
 RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ngl227;
 RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.R.,
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 RA Cerdeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
 RA Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,
 RA Rabbिनowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
 RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
 RT "Comparative analysis of the genome sequences of Bordetella pertussis,
 RT Bordetella parapertussis and Bordetella bronchiseptica.";
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL; BX640412; CAE44677.1; -.
 DR GO; GO:0005381; F:iron ion transporter activity; IEA.
 DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
 DR InterPro; IPR002491; Peripla_BP.
 DR Pfam; PF01497; Peripla_BP_2; 1.
 KW Complete proteome.
 SQ SEQUENCE 289 AA; 29505 MW; 3B80C28C1D8940AD CRC64;

 Query Match 100.0%; Score 33; DB 2; Length 289;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 LGTGPR 6
 |||||
 Db 261 LGTGPR 266

RESULT 7
 Q8P6Z3

ID Q8P6Z3 PRELIMINARY; PRT; 491 AA.
 AC Q8P6Z3;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Protein-glutamate methylesterase.
 GN Name=cheB; OrderedLocusNames=XCC2822;
 OS Xanthomonas campestris (pv. campestris).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 OX NCBI_TaxID=340;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 33913 / NCPPB 528;
 RX MEDLINE=22022145; PubMed=12024217; DOI=10.1038/417459a;
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
 RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
 RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
 RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
 RA Setubal J.C., Kitajima J.P.;
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing
 RT host specificities.";
 RL Nature 417:459-463(2002).
 DR EMBL; AE012394; AAM42094.1; -.
 DR HSSP; P04042; 1CHD.
 DR GO; GO:0004871; F:signal transducer activity; IEA.
 DR GO; GO:0006935; P:chemotaxis; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR GO; GO:0007165; P:signal transduction; IEA.
 DR InterPro; IPR003313; AraC_binding.
 DR InterPro; IPR000673; CheB_methylest.
 DR InterPro; IPR011247; Chmtx_methlestr.
 DR InterPro; IPR003006; Ig_MHC.
 DR Pfam; PF01339; CheB_methylest; 1.
 DR PIRSF; PIRSF036461; Chmtx_methlestr; 1.
 DR ProDom; PD005328; CheB_methylest; 1.
 DR PROSITE; PS50122; CHEB; 1.
 DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 491 AA; 51780 MW; 379E3413A027F619 CRC64;

Query Match 100.0%; Score 33; DB 2; Length 491;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
 |||||
 Db 248 LGTGPR 253

RESULT 8

Q8IOF1

ID Q8IOF1 PRELIMINARY; PRT; 563 AA.
 AC Q8IOF1;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE CG31538-PA (AT27831p).
 GN ORFNames=CG31538;

OS *Drosophila melanogaster* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; *Drosophila*.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
 RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426065; PubMed=12537568;
 RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
 RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
 RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
 RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
 RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
 RT "Finishing a whole-genome shotgun: Release 3 of the *Drosophila*
 RT *melanogaster* euchromatic genome sequence.";

RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426070; PubMed=12537573;
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 RA Ashburner M., Celniker S.E.;
 RT "The transposable elements of the Drosophila melanogaster euchromatin:
 RT a genomics perspective.";
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a
 RT systematic review.";
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
 RN [5]
 RP SEQUENCE FROM N.A.
 RG FlyBase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RG FlyBase;
 RL Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
 RA Patel S., Phouanenvong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celniker S.;
 RL Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AE003603; AAN13275.1; -.
 DR EMBL; BT001356; AAN71111.1; -.
 DR FlyBase; FBgn0051538; CG31538.
 SQ SEQUENCE 563 AA; 63800 MW; 8E054274E710C583 CRC64;

Query Match 100.0%; Score 33; DB 2; Length 563;
 Best Local Similarity 100.0%; Pred. No. 3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
 |||||
 Db 269 LGTGPR 274

RESULT 9

Q8NEG7

ID Q8NEG7 PRELIMINARY; PRT; 585 AA.

AC Q8NEG7;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Similar to mouse 1700027J05Rik protein.
 GN Name=MGC33692;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA Strausberg R.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; BC031069; AAH31069.1; -.
 SQ SEQUENCE 585 AA; 66464 MW; 2B2D5F46647D448C CRC64;

Query Match 100.0%; Score 33; DB 2; Length 585;
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
 |||||
 Db 5 LGTGPR 10

RESULT 10

Q6YBW4

ID Q6YBW4 PRELIMINARY; PRT; 736 AA.
 AC Q6YBW4;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE TACC3.
 GN Name=TACC3;
 OS *Oryctolagus cuniculus* (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; *Oryctolagus*.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=15207008;
 RA Still I.H., Vettaikkorumakankauv A.K., DiMatteo A., Liang P.;
 RT "Structure-function evolution of the Transforming acidic coiled coil
 RT genes revealed by analysis of phylogenetically diverse organisms.";
 RL BMC Evol. Biol. 4:16-16(2004).
 DR EMBL; AY161270; AAO25635.1; -.
 DR InterPro; IPR007707; TACC.
 DR Pfam; PF05010; TACC; 1.
 SQ SEQUENCE 736 AA; 77061 MW; A798FB1C177EF3C8 CRC64;

Query Match 100.0%; Score 33; DB 2; Length 736;
 Best Local Similarity 100.0%; Pred. No. 3.9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
 |||||
 Db 207 LGTGPR 212

RESULT 11

Q6ZRS4

ID Q6ZRS4 PRELIMINARY; PRT; 896 AA.
 AC Q6ZRS4;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein FLJ46145.
 OS *Homo sapiens* (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; *Homo*.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA Suzuki O., Sasaki N., Aotsuka S., Shoji T., Ichihara T., Shiohata N.,
 RA Matsumoto K., Hirano M., Sano S., Nomura R., Yoshikawa Y.,
 RA Matsumura Y., Moriya S., Chiba E., Momiyama H., Onogawa S.,
 RA Kaeriyama S., Satoh N., Matsunawa H., Takahashi E., Kataoka R.,
 RA Kuga N., Kuroda A., Satoh I., Kamata K., Takami S., Terashima Y.,
 RA Watanabe M., Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A.,
 RA Ishii S., Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
 RA Kimura K., Yamashita H., Matsuo K., Nakamura Y., Sekine M.,
 RA Kikuchi H., Kanda K., Wagatsuma M., Murakawa K., Kanehori K.,
 RA Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B., Suzuki Y.,
 RA Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
 RL Submitted (JUL-2003) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AK128026; BAC87235.1; -.
 SQ SEQUENCE 896 AA; 98946 MW; DBC9EF0E6CF7B2C0 CRC64;

Query Match 100.0%; Score 33; DB 2; Length 896;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
|||||
Db 578 LGTGPR 583

RESULT 12

Q9P5K6

ID Q9P5K6 PRELIMINARY; PRT; 946 AA.
AC Q9P5K6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Related to glucan 1, 4-alpha-glucosidase.
GN Name=B23L21.230;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AL356172; CAB91691.2; -.
DR PIR; T49710; T49710.
SQ SEQUENCE 946 AA; 101461 MW; A8564328338B6E1C CRC64;

Query Match 100.0%; Score 33; DB 2; Length 946;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
|||||
Db 807 LGTGPR 812

RESULT 13

Q8KLS0

ID Q8KLS0 PRELIMINARY; PRT; 1095 AA.
AC Q8KLS0;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative histidine protein kinase.
GN Name=cheA3;
OS Rhodobacter sphaeroides (Rhodopseudomonas sphaeroides).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC Rhodobacteraceae; Rhodobacter.

OX NCBI_TaxID=1063;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WS8N;
 RA Porter S.L., Warren A.V., Armitage J.P.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AJ488585; CAD32761.1; -.
 DR GO; GO:0005622; C:intracellular; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
 DR GO; GO:0006935; P:chemotaxis; IEA.
 DR GO; GO:0000160; P:two-component signal transduction system (p. . .; IEA.
 DR InterPro; IPR003594; ATPbind_ATPase.
 DR InterPro; IPR002545; CheW.
 DR InterPro; IPR008207; Hpt.
 DR InterPro; IPR008208; Hpt_N.
 DR Pfam; PF01584; CheW; 1.
 DR Pfam; PF02518; HATPase_c; 1.
 DR Pfam; PF01627; Hpt; 1.
 DR ProDom; PD003142; Hpt_N; 1.
 DR SMART; SM00260; CheW; 1.
 DR SMART; SM00073; HPT; 1.
 DR PROSITE; PS50894; HPT; 1.
 KW Kinase.
 SQ SEQUENCE 1095 AA; 114521 MW; F43CF5A0EB4F3F0E CRC64;

Query Match 100.0%; Score 33; DB 2; Length 1095;
 Best Local Similarity 100.0%; Pred. No. 5.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
 |||||
 Db 1031 LGTGPR 1036

RESULT 14
 O15069

ID O15069 PRELIMINARY; PRT; 1522 AA.
 AC O15069;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE KIAA0363 protein (Fragment).
 GN Name=KIAA0363;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=97349984; PubMed=9205841;
 RA Nagase T., Ishikawa K., Nakajima D., Ohira M., Seki N., Miyajima N.,
 RA Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. VII.
 RT The complete sequences of 100 new cDNA clones from brain which can

RT code for large proteins in vitro."
 RL DNA Res. 4:141-150(1997).
 DR EMBL; AB002361; BAA20818.1; -.
 DR InterPro; IPR002715; NAC.
 DR Pfam; PF01849; NAC; 1.
 FT NON_TER 1 1
 SQ SEQUENCE 1522 AA; 156998 MW; 5779025D6AB66C04 CRC64;

Query Match 100.0%; Score 33; DB 2; Length 1522;
 Best Local Similarity 100.0%; Pred. No. 7.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
 |||||
 Db 1251 LGTGPR 1256

RESULT 15

Q8Y9T8

ID Q8Y9T8 PRELIMINARY; PRT; 2013 AA.
 AC Q8Y9T8;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Putative peptidoglycan bound protein (LPXTG motif).
 GN OrderedLocusNames=lmo0435;
 OS Listeria monocytogenes.
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
 OX NCBI_TaxID=1639;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EGD-e / Serovar 1/2a;
 RX MEDLINE=21537279; PubMed=11679669; DOI=10.1126/science.1063447;
 RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
 RA Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
 RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
 RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
 RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
 RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
 RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
 RA Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
 RA Rammel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
 RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
 RT "Comparative genomics of Listeria species."
 RL Science 294:849-852(2001).
 CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
 CC an amide bond (By similarity).
 DR EMBL; AL591975; CAC98514.1; -.
 DR PIR; AD1129; AD1129.
 DR ListiList; LMO0435; -.
 DR GO; GO:0009986; C:cell surface; IEA.
 DR GO; GO:0005618; C:cell wall; IEA.
 DR Pfam; PF00746; Gram_pos_anchor; 1.
 DR SMART; SM00089; PKD; 6.
 DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
 DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.

KW Cell wall; Complete proteome; Peptidoglycan-anchor.
SQ SEQUENCE 2013 AA; 219294 MW; 0D8A79F9EC659A90 CRC64;

Query Match 100.0%; Score 33; DB 2; Length 2013;
Best Local Similarity 100.0%; Pred. No. 9.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LGTGPR 6
|||||
Db 892 LGTGPR 897

Search completed: February 11, 2005, 18:22:47
Job time : 9.23616 secs